

REMARKS

This Amendment is in response to the Office Action of November 27, 2007 in which all of the pending claims 1-16 were rejected. New claims 17-20 are added. Independent device claim 17 is similar to amended independent method claim 1. No new claim fees are due. This amendment is timely and no extension of time is needed. However if these assumptions are wrong, the Commissioner is authorized to deduct the appropriate fee from our deposit account number 23-0442 and if a petition is needed, to consider this paper as the required petition.

I. Formal Objections/Rejections

Regarding the statutory subject matter rejection of claims 7, 8 and 9, claims 7 and 8 have been cancelled and claim 9 amended so as to be directed to a computer readable medium. Withdrawal of the 35 USC 101 rejection is requested.

II. Claimed Invention

The subject matter of method claim 1, as amended, is as follows:

- A. receiving an initial user input causing a mobile communication device to be prepared for receiving an acoustic input of the user to perform speech recognition thereon;
- B1. receiving said acoustic input of the user and
- B2. performing speech recognition thereon;
- C. performing a back-up operation to enable said user to provide manual input in case of failure of said speech recognition of said acoustic input as follows;
- C1a. upon receiving a first manual user input by a multiple switching component, which is capable to exhibit a first input value and a second input value,

- C1b. displaying a list of a first set of data records or displaying a second set of data records in accordance with said first input value and said second input value of said first manual user input; and
- C2a. upon receiving a second manual user input identifying one data record of said displayed first set of data records or of said second set of data records,
- C2b. transmitting an instruction corresponding to said identified data record to at least one application of a plurality of applications executable on said mobile communication device.

### III. Prior Art

The Examiner raises a 35 USC 103 rejection against the pending claims 1 to 16 as being unpatentably obvious over *Gerson* (US 6,868,385) in view of *Waibel et al* (US 5,855,000).

*Gerson* in general relates to a wireless communication environment, in which at least one subscriber unit (mobile unit) is in wireless communication with an infrastructure. The subscriber unit implements a speech recognition client and the infrastructure comprises a speech recognition server. The subscriber unit takes as input an unencoded speech signal that is subsequently parameterized by the speech recognition client. Voice and data processing operations of the subscriber unit are shown in Figures 3 and 7. The parameterized speech is provided to the speech recognition server, which performs speech recognition analysis on the parameterized speech. Voice and data processing operations of the server are shown in Figures 5 and 6. As part of a client-server speech recognition arrangement, the speech recognition analyzer 504 takes speech recognition parameter vectors from a subscriber unit and completes recognition processing (see column 11, lines 51-4). In Figure 6, information signals, based in part upon any recognized utterance identified by the speech recognition analysis, are provided to the subscriber unit. The information signals may comprise control signals used to control the subscriber

unit itself or to control one or more devices coupled to the subscriber unit. The information signals may comprise control signals used to control the subscriber unit itself or to control one or more devices coupled to the subscriber unit. Such data signals can be used to locally develop control signals or may lead to the provision of additional user data to the speech recognition server which can respond with additional information signals as described above (cf. col. 1, line 66 to col. 2, line 27 of *Gerson*).

A detailed embodiment of operation of the subscriber unit of *Gerson* is described in col. 13, line 57 to col. 14, line 29, which refers to the flow chart shown in Fig. 7. In this embodiment, the unencoded speech signal received by the subscriber unit is converted and analyzed to provide parameterized speech signal, which is processed at the speech recognition server. The speech recognition server may use lookup-tables, pattern matching and/or similar mechanisms to correlate a specific recognized utterance or string of utterances to one or more predefined information signals (cf. col. 13, lines 1 to 20). The subscriber unit receives the information signals and the information signals are operated upon by or used to control operations of the subscriber unit itself or any devices coupled to the subscriber unit. For instance, when the information signals comprise data, the data can be used to locally generate (at the subscriber unit) control signals. For instance, the receipt of a phone number can be used to trigger a control signal instructing the subscriber unit to dial the phone number. There is also a scenario mentioned at column 13, lines 40-48 where a user originally requests a party's phone number by name. If, however, ambiguity exists because of multiple parties having the same name, the information signals provided in response may request the user to select one of the parties through use of a touch-tone pad (i.e., using DTMF tones) or by responding to the name of one of the parties.

Further, *Gerson* describes a local speech recognizer 303, which may receive parameter vectors from a speech recognition front-end 302, to perform speech recognition analysis thereon. The recognized utterances may be passed to an application 307 including a detector application, which for example compares the

recognized utterances against a list of predetermined utterances (cf. col. 9, line 62 to col. 10, line 10).

*Waibel et al* in general relates to a methodology for locating an error within a recognition hypothesis. *Waibel et al* addresses the problem of repairing the output of a recognition engine. An incorrectly recognized primary input signal produced by a (speech) recognition hypothesis is enabled to be corrected through a secondary input signal, which is independent of the primary input signal. This means the primary and secondary input signals may be of different modalities (e.g. speech and nonverbal spelling, speech and verbal spelling, speech and writing etc.) (cf. col. 3, lines 20 to 28 and lines 46 to 49).

With reference to Fig. 1 *Waibel et al* describes a speech recognition engine 14, which receives audio input referred to as a primary utterance through a microphone and input electronics. The output of the speech recognition engine 14 is input to the correction and repair module 12.

The output of the correction and repair module 12 may be displayed on a screen and the correction and repair module 12 may be responsive to correction input through a keyboard, use of a mouse or other pointing devices to highlight words on screen. Errors are highlighted by one of the aforementioned means (cf. col. 4, line 65 to col. 5, line 4; col. 5, lines 36 to 45; and col. 5, line 53 to col. 6, line 4).

On the basis of the primary utterance received by the speech recognition engine 14 a recognition hypothesis is produced. The recognition hypothesis is displayed and some error can be identified in the displayed recognition by the user using the aforementioned highlighting. If error are identified, the hypothesis is rejected and a secondary input signal is received from the user and based on the type of user input (corresponding to the spelling recognition engine 16, cursive handwriting recognition engine 18, gesture recognition engine 19) a repair mode is selected (cf. col. 6, line 51 to col. 7, line 2).

As illustratively described in col. 8, lines 17 to lines 25, a user may speak “I would like to meet Monday around noon” while the speech recognition engine provides a recognition hypothesis of “I would like to meet one day around noon”. The user rejects the hypothesis by identifying the word “one day” as being in error by touching the touch sensitive screen.” The user then provides a secondary input signal, which in this example is a respeaking of the located error.

#### IV. Applicant's comments

In view of the above amended set of claims which are based on the amendments previously filed in the parallel European patent application of this family and in light of the comments which follow, the applicant respectfully requests reconsideration.

The amended set of claims includes a new claim 1 which is formed on the basis of the pending claim 1 and including the feature that a manual user input is a backup operation in case of failure of the speech recognition (cf. paragraph [0045] of the US application publication corresponding to the paragraph beginning at page 8, line 20 of the application) and that upon receiving a first manual user input by a multiple switching component there is displayed a list of a first set of data records or a second set of data records, and upon receiving a second manual user input identifying one data record of the displayed first or second sets of data records, an instruction corresponding to the identified data record is transmitted to at least one application of a plurality of applications executable on the mobile communication device.

The Examiner states that *Waibel et al* describe the displaying of an n-best list and refers to col. 5, lines 63-67 and col. 9, lines 55-67 thereof. Respectfully, as the n-best list does not correspond to a list displayed to a user, this interpretation of *Waibel et al* is incorrect. *Waibel et al* teach the use of n-best lists, i.e., primary n-best list and secondary n-best list, as inputs to the correction and repair module 12 (cf. col. 10, lines 11 to 20 and lines 48 to 56). Each n-best list includes result list

outputted by the corresponding recognition engine, wherein each result list includes the n-best recognition results rated by score values.

The correction and repair module 12 processes the inputted n-best lists and outputs the top-choice resulting from the processing of the score values of the n-best lists to the user. This means, the correction and repair module 12 does not output a list but only the top-choice, i.e. the choice having the highest total score value, resulting from both n-best lists. Only the output of the correction and repair module 12 is shown to the user on the screen of the computer (cf. col. 5, lines 63 to 64 of *Gerson*) but not any of the n-best lists as stated by the Examiner.

On the first hand, the lists are defined in detail, namely, the first set of data records (i.e. corresponding to a first list) represents telephone directory entries associated with voice tags and selectable by speech recognition and the second set of data records (i.e. corresponding to a second list) represents device functions controllable by speech recognition and associated with voice tags. According to the step of “displaying a list of the first or second set of data records” either a first list comprising list entries of the data records representing telephone directory entries associated with voice tags and selectable by speech recognition and the second set of data records or a second list comprising list entries of the data records representing device functions controllable by speech recognition and associated with voice tags is displayed. Neither the displaying of a list nor the separation into the both above defined lists is described or suggested in any of the cited prior art documents.

Moreover, even when considering the n-best lists inputted to the correction and repair module 12 described by *Waibel et al*, the n-best lists thereof are processed when an error is identified by the user. *Waibel et al* teach in col. 9, lines 56 to 66 that the n-best list is used by the correction and repair module 12 in the context of a speech recognition engine that produces an n-best list for unimodal repair with correlation used by the correction and repair module 12. Accordingly, a secondary input namely a respeaking by the user or information inputted in some other manner

such as spelling, writing, etc. of the erroneous subsection of the primary utterance is required for correlation processing (cf. also col. 10, lines 57 to 65).

The subject matter of the amended claim 1 defines that a back-up operation (feature C) is provided for enabling the user to provide manual input in two ways in case of failure of the speech recognition. Upon receiving a first manual user input (feature C1a) displaying a list of a first set of data records or displaying a second set of data records in accordance with the value of the first manual user input (C1b). Upon receiving a second manual user input identifying one data record of the displayed first set of data records or of the second set of data records (feature C2a), transmitting (feature C2b) an instruction corresponding to the identified data record to at least one application of a plurality of applications executable on the mobile communication device.

Neither reference shows or suggests the performance of a backup operation as claimed with a first or second set of data records displayed depending on user manual input followed by a second user input to transmit an instruction.

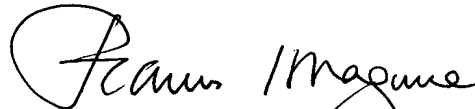
Moreover, the separation into the aforementioned two lists improves significantly the usability of the manual backup operation as an alternative of the acoustic input and speech recognition. Upon first manual user actuation, the user may select one of the list and upon second manual user actuation, the user may select a list entry of the respective selected list, wherein the first list includes the telephone directory entries associated with voice tags and selectable by speech recognition and the second list includes the device functions controllable by speech recognition and associated with voice tags. Hence, a fast and feasible access to a user desired list entry is ensured by the categorized compositions of the lists.

Withdrawal of the obviousness rejection of claims 1-16 is requested. Claim 17 is similar to claim 1 and is allowable for the same reasons as given above. Independent device claim 10 is more specific concerning the devices used to actuate user input and although a back-up procedure is not explicitly mentioned, the sub-features C1a-C2b are essentially the same.

V. Conclusion

The objections and rejections of the Office Action of November 27, 2007, having been obviated by amendment or shown to be inapplicable, withdrawal thereof is requested and passage of claims 1-20 to issue is earnestly solicited.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Francis J. Maguire". The signature is written in a cursive, flowing style with a large initial "F" and "M".

Francis J. Maguire  
Attorney for the Applicant  
Registration No. 31,391

FJM/mo  
WARE, FRESSOLA, VAN DER SLUYS  
& ADOLPHSON LLP  
755 Main Street, P.O. Box 224  
Monroe, Connecticut 06468  
(203) 261-1234